Adaptive Predictive Coding of Speech Signals

By B. S. ATAL and M. R. SCHROEDER

(Manuscript received December 13, 1968)

We describe in this paper a method for efficient encoding of speech signals, based on predictive coding. In this coding method, both the transmitter and the receiver estimate the signal's current value by linear prediction on the previously transmitted signal. The difference between this estimate and the true value of the signal is quantized, coded and transmitted to the receiver. At the receiver, the decoded difference signal is added to the predicted signal to reproduce the input speech signal. Because of the nonstationary nature of the speech signals, an adaptive linear predictor is used, which is readjusted periodically to minimize the mean-square error between the predicted and the true value of the signals.

The predictive coding system was simulated on a digital computer. The predictor parameters, comprising one delay and nine other coefficients related to the signal spectrum, were readjusted every 5 milliseconds. The speech signal was sampled at a rate of 6.67 kHz, and the difference signal was quantized by a two-level quantizer with variable step size. Subjective comparisons with speech from a logarithmic PCM encoder (log-PCM) indicate that the quality of the synthesized speech signal from the predictive coding system is approximately equal to that of log-PCM speech encoded at 6 *bits/sample.*

Preliminary studies suggest that the binary difference signal and the predictor parameters together can be transmitted at approximately 10 kilobits/second which is several times less than the bit rate required for log-PCM encoding with comparable speech quality.

I. INTRODUCTION

The aim of efficient coding methods[1] is to reduce the channel capacity required to transmit a signal with specified fidelity. To achieve this objective, it is often essential to reduce the redundancy of the transmitted signal. One well-known procedure for reducing signal redundancy

is predictive coding.*[2-5] In predictive coding, redundancy is reduced by subtracting from the signal that part which can be predicted from its past. For many signals, the first-order entropy of the difference signal is much smaller than the first-order entropy of the original signal; thus, the difference signal is better suited to memoryless encoding than the original signal. Predictive coding offers a practical way of coding signals efficiently without requiring large codebook memories.

Many previous speech coding methods[6] have employed schemes which attempt to separate the contributions of the vocal excitation from that of the vocal-tract transmission function. The well-known channel vocoder of Dudley[7] was the first attempt in this direction. Although vocoders can reproduce intelligible speech, there is appreciable loss in naturalness and speech quality. This degradation in speech quality arises from various operations in the vocoding process, which are either inaccurately performed or are based on certain idealized approximations of speech production and perception processes.

The present paper describes a different approach[8,9] to encoding of speech signals, based on predictive coding, which avoids the difficulties encountered in vocoders and vocoder-like devices. Although predictive coding utilizes such well-known characteristics of speech signals as pitch and formant structure, its operation does not rely solely upon a rigid parameterization of the speech signal. That part of the speech signal which cannot be represented in terms of these characteristics is not discarded but suitably encoded and transmitted to the receiver where it is used in the synthesis of a close replica of the original speech waveform.

Previous studies of predictive coding systems for speech signals[10] have been limited to linear predictors with fixed coefficients. However, due to the nonstationary nature of the speech signals, a fixed predictor cannot predict the signal values efficiently at all times. For example, the speech waveform is approximately periodic during voiced portions; thus, a good prediction of the present value of the signal can be based on the value of the signal exactly one period earlier. However, the period of the speech signal varies with time. The predictor, therefore, must change with the changing period of the input speech signal. In the predictive coding system described below, the linear predictor is adaptive; it is readjusted periodically to match the time-varying characteristics of the input speech signal. The parameters of the linear predictor are optimized to obtain an efficient prediction in the sense that

* Another name often used for this kind of encoding is Differential Pulse Code Modulation.

the mean-square error between the predicted value and the true value of the signal is minimum.

II. PREDICTIVE CODING SYSTEM

2.1 *Description*

A block diagram illustrating the principle of predictive coding is shown in Fig. 1. The input signal $s(t)$ is sampled at the Nyquist rate to produce the samples s_n of the signal. The predictor forms an estimate $\hat{s}_n$ of the signal's present value based on the past samples r_{n-1}, r_{n-2}, $\cdots$ of the reconstructed signal at the transmitter. The predicted value $\hat{s}_n$ of the signal is next subtracted from the signal value s_n to form the difference δ_n, which is quantized, encoded, and transmitted to the receiver. At the same time, the transmitted signal is decoded at the transmitter and the signal reconstructed in exactly the same manner as is done at the receiver. The reconstructed signal is then used to predict the next sample of the input signal.

At the receiver, the transmitted signal is decoded and added to the predicted value of the signal to form the samples r'_n of the reconstructed signal. The predictor used at the receiver is identical to one employed at the transmitter. The samples r'_n of the reconstructed signal are finally low-pass filtered to produce the output signal $r'(t)$.

2.2 *Signal-to-Quantizing Noise Ratio*

Consider the predictive coding system shown in Fig. 1. Let P_s be the mean-square value of the input signal samples s_n, P_δ be the mean-square value of the difference signal samples δ_n, P_q be the mean-square value of the quantizing noise in the decoded difference signal δ'_n, and

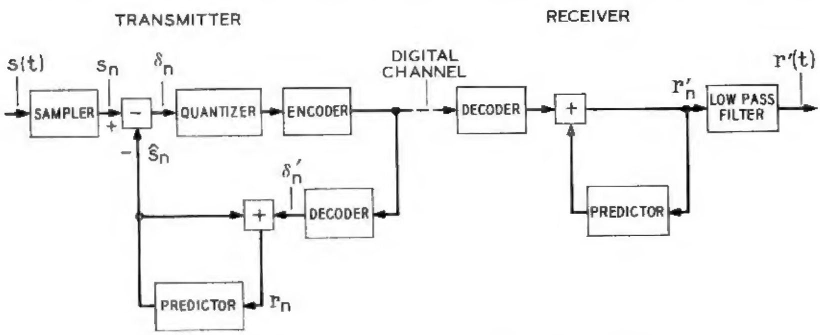

Fig. 1—Block diagram of a predictive coding system.

P_e be the mean-square value of the quantizing noise in the reconstructed signal r'_n. We will now show that, in the absence of digital channel transmission errors, the signal-to-quantizing noise ratio P_s/P_e of the reconstructed signal is given by

$$\frac{P_s}{P_e} = \frac{P_s}{P_\delta} \cdot \frac{P_\delta}{P_q}. \tag{1}$$

In other words, the signal-to-quantizing noise ratio of the reconstructed signal *exceeds* the signal-to-quantizing noise ratio of the decoded difference signal by a factor equal to the ratio of the mean-square value of the input signal to the mean-square value of the difference signal. The predictive coding system is thus superior to a straight PCM system whenever P_s/P_δ is much greater than 1. For a signal such as speech, this is indeed true. The results obtained by computer simulation of the predictive coding system (see Section 3.3) show that P_s/P_δ is about 100 for speech signals. By using predictive coding, one could thus expect improvement of about 20 dB in signal-to-quantizing noise ratio over a PCM system using identical quantizing levels.

To prove equation (1), we will first show that the error between any sample of the reconstructed signal and the corresponding sample of the input signal is identical to the error introduced by the quantizer, the encoder and the decoder.

The error e_n between the sample r'_n of the reconstructed signal and the sample s_n of the input signal is given by

$$e_n = r'_n - s_n . \tag{2}$$

In the absence of digital channel transmission errors, we can replace r'_n in equation (2) by r_n and rewrite equation (2) as

$$e_n = (r_n - \hat{s}_n) - (s_n - \hat{s}_n). \tag{3}$$

It is readily seen in Fig. 1 that

$$r_n = \delta'_n + \hat{s}_n \tag{4}$$

and

$$\delta_n = s_n - \hat{s}_n . \tag{5}$$

On combining equations (3), (4) and (5), one obtains

$$e_n = \delta'_n - \delta_n . \tag{6}$$

The right side of equation (6) represents the error introduced by the quantizer, the encoder, and the decoder. Thus, the error in the nth

sample of the reconstructed signal is identical to the error in the nth sample of the decoded difference signal.

The signal-to-quantizing noise ratio of the reconstructed signal is by definition P_s/P_e and can be written as

$$\frac{P_s}{P_e} = \frac{P_s}{P_\delta} \cdot \frac{P_\delta}{P_e}. \tag{7}$$

Since the mean-square value P_e of the quantizing noise in the reconstructed signal is identical to the mean-square value P_q of the quantizing noise in the decoded difference signal, P_e on the right side of equation (7) can be replaced by P_q , and one obtains

$$\frac{P_s}{P_e} = \frac{P_s}{P_\delta} \cdot \frac{P_\delta}{P_q}. \tag{1}$$

III. APPLICATION OF PREDICTIVE CODING TO SPEECH SIGNALS

3.1 *Linear Prediction of Speech Signals*

Two of the main causes of redundancy in speech are:

(*i*) Quasi-periodicity during voiced segments[6] and,
(*ii*) Lack of flatness of the short-time spectral envelope.[6]

The exact form of the predictor for the speech wave depends on the model used to represent the human speech production process. A reasonable model for the production of voiced speech sounds is obtained by representing them as the output of a discrete linear time-varying filter which is excited by a quasi-periodic pulse train (see Fig. 2). The output of the linear filter at any sampling instant is a linear combination of the past p output samples and the input. The number of past samples p is given by twice the number of resonances (formants) of the vocal tract which are contained in the frequency range of interest. For example, in the case of speech signals band-limited to 3 kHz, it can be

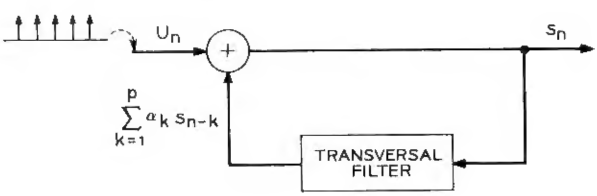

Fig. 2—Model for the production of voiced speech sounds.

assumed that there are typically three to four formants.[6] A suitable value of p is thus 8.

Let s_n and U_n be the amplitudes of the output and input signals (see Fig. 2) at the nth sampling instant. The nth output sample s_n is then given by

$$s_n = \sum_{k=1}^{p} \alpha_k s_{n-k} + U_n , \tag{8}$$

where

$$U_n = \beta U_{n-M} , \tag{9}$$

M is the period of the excitation signal and β takes account of the variation of the amplitude of the input pulse train from one period to the next. For natural speaking conditions, the period of the excitation signal is usually below 15 milliseconds, and, as a first approximation, the effect of time variation of the coefficients α_k from one pitch period to the next can be neglected. Under this assumption, we find

$$s_n - \beta s_{n-M} = \sum_{k=1}^{p} \alpha_k (s_{n-k} - \beta s_{n-k-M}) + U_n - \beta U_{n-M} . \tag{10}$$

Since $U_n = \beta U_{n-M}$, equation (10) reduces to

$$s_n = \beta s_{n-M} + \sum_{k=1}^{p} \alpha_k (s_{n-k} - \beta s_{n-k-M}), \tag{11}$$

which determines completely the structure of the linear predictor.

A block diagram of the predictor as described by equation (11) is shown in Fig. 3. The delay M as well as the parameters $\alpha_1 , \alpha_2 , \cdots , \alpha_p$

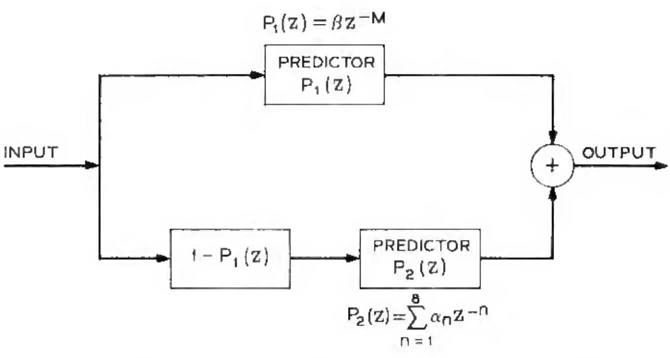

Fig. 3—Block diagram of the predictor for speech signals.

and β are variable and are readjusted periodically to match the characteristics of the input speech signal. Ideally the readjustment of the predictor parameters need be done only when there are significant changes in the characteristics of the speech signal. This implies that the predictor should be readjusted at short intervals during transitions and at long intervals during steady state portions of the speech signal and, consequently, a long buffer storage is needed to ensure transmission of parameters at a uniform rate on the channel. In order to avoid the use of a long buffer storage, the predictor parameters were readjusted at a fixed time interval in our study. This time interval was chosen to be 5 milliseconds to ensure that the prediction be efficient even during rapidly changing segments of the speech wave.

For unvoiced sounds, the quasi-periodic excitation U_n in equation (8) is replaced by a noise-like excitation. Generally speaking, the transfer function of the filter for unvoiced sounds must include poles as well as zeros. However, we find that for all practical purposes it is sufficient to include only the effect of poles. Equation (11), thus, represents the linear predictor for unvoiced sounds too if β is assumed zero.

3.2 *Determination of Predictor Parameters*

The predictor parameters are determined by minimizing the mean-square error between the actual speech sample and its predicted value. The predicted value $\hat{s}_n$ of the nth speech sample is given by

$$\hat{s}_n = \beta s_{n-M} + \sum_{k=1}^{p} \alpha_k (s_{n-k} - \beta s_{n-k-M}). \tag{12}$$

The prediction error sample E_n is then given by

$$\begin{aligned} E_n &= s_n - \hat{s}_n \\ &= (s_n - \beta s_{n-M}) - \sum_{k=1}^{p} \alpha_k (s_{n-k} - \beta s_{n-k-M}). \end{aligned} \tag{13}$$

The mean-square prediction error $\langle E_n^2 \rangle_{\text{av}}$ is given by

$$\langle E_n^2 \rangle_{\text{av}} = \frac{1}{N} \sum_n E_n^2 , \tag{14}$$

where the sum extends over all the samples in the time interval during which the predictor is to be optimum.

The problem of minimizing the mean-square error $\langle E_n^2 \rangle_{\text{av}}$ by suitable selection of the predictor parameters does not admit a straightforward solution due to the presence of the delay parameter M in equation (13).

A sub-optimum solution was obtained by minimizing the total error in two steps. First the parameters β and M are determined such that the error E_1, defined by

$$E_1 = \frac{1}{N} \sum_n (s_n - \beta s_{n-M})^2 = \langle (s_n - \beta s_{n-M})^2 \rangle_{\text{av}} , \tag{15}$$

is minimum. Using these values of β and M, the mean-square error $\langle E_n^2 \rangle_{\text{av}}$ is minimized by a suitable choice of parameters $\alpha_1, \cdots, \alpha_p$.

To find the values of the parameters β and M which minimize the error E_1 as defined in equation (15), we first set the partial derivative of E_1 with respect to β equal to zero:

$$\begin{aligned} \frac{\partial E_1}{\partial \beta} &= -2\langle (s_n - \beta s_{n-M}) s_{n-M} \rangle_{\text{av}} \\ &= 0, \end{aligned} \tag{16}$$

where the $\langle\ \rangle_{\text{av}}$ indicates the averaging over all the samples in the given 5-millisecond time segment during which the predictor is to be optimum.

On solving for β from equation (16), we obtain

$$\beta = \langle s_n s_{n-M} \rangle_{\text{av}} / \langle s_{n-M}^2 \rangle_{\text{av}} . \tag{17}$$

We next substitute the value of β from equation (17) into equation (15). After rearrangement of terms, we obtain

$$E_1 = \langle s_n^2 \rangle - \langle s_n s_{n-M} \rangle_{\text{av}}^2 / \langle s_{n-M}^2 \rangle_{\text{av}} . \tag{18}$$

Since the first term on the right side of equation (18) does not depend on M, it can be omitted in finding the minimum value of the error. Further, E_1 is minimum if the second term on the right side of equation (18) is maximum. The optimum value of M is thus determined from the location of the maximum of the normalized correlation coefficient ρ given by

$$\rho = \{ \langle s_n s_{n-M} \rangle_{\text{av}} \} / \{ \langle s_n^2 \rangle_{\text{av}} \langle s_{n-M}^2 \rangle_{\text{av}} \}^{\frac{1}{2}}, \qquad M > 0. \tag{19}$$

Next, the predictor parameters $\alpha_1, \cdots, \alpha_p$ are obtained such that the mean-square error $\langle E_n^2 \rangle_{\text{av}}$ as given in equation (14) with β and M fixed at their optimum values is minimum. Let

$$v_n = s_n - \beta s_{n-M} . \tag{20}$$

The error $\langle E_n^2 \rangle_{\text{av}}$ is then given by

$$\langle E_n^2 \rangle_{\text{av}} = \left\langle \left\{ v_n - \sum_{k=1}^{p} \alpha_k v_{n-k} \right\}^2 \right\rangle_{\text{av}} . \tag{21}$$

The optimum values of the coefficients $\alpha_1, \cdots, \alpha_p$ which minimize $\langle E_n^2 \rangle_{av}$ are obtained by setting the partial derivatives of $\langle E_n^2 \rangle_{av}$ with respect to $\alpha_1, \cdots, \alpha_p$ equal to zero. Or,

$$\frac{\partial \langle E_n^2 \rangle_{av}}{\partial \alpha_j} = \left\langle \left(v_n - \sum_{k=1}^{p} \alpha_k v_{n-k} \right) v_{n-j} \right\rangle_{av}, \\ = 0 \quad \text{for} \quad j = 1, 2, \cdots, p. \tag{22}$$

Equation (22) can be rewritten in matrix notation as

$$\Phi \mathbf{a} = \boldsymbol{\psi}, \tag{23}$$

where Φ is a p by p matrix with its (ij)th term φ_{ij} given by

$$\varphi_{ij} = \langle v_{n-i} v_{n-j} \rangle_{av}, \tag{24}$$

$\mathbf{a}$ is a p-dimensional vector whose jth component is a_j and $\boldsymbol{\psi}$ is a p-dimensional vector whose jth component ψ_j is given by

$$\psi_j = \langle v_n v_{n-j} \rangle_{av}. \tag{25}$$

The optimum predictor coefficients $\alpha_1, \alpha_2, \cdots, \alpha_p$ are obtained by solving equation (23) for $\mathbf{a}$. For the case when Φ is a nonsingular matrix, the solution of equation (23) presents no difficulty. The vector $\mathbf{a}$ can be obtained by multiplying $\boldsymbol{\psi}$ with the inverse of the matrix Φ. A more efficient computational procedure[11] for solving equation (23), which does not involve matrix inversion, takes advantage of the fact that Φ is a symmetric matrix, and thus can be expressed as the product of a triangular matrix and its transpose. Equation (23) can then be written as three separate matrix equations. These equations involve triangular matrices only and their solutions can be expressed by a set of recursive equations.[11]

A singular Φ matrix implies that one or more of its eigenvalues is zero. The matrix Φ can be modified to become nonsingular by adding a small positive constant to its diagonal elements. Equation (23) is solved again with the matrix Φ replaced by the matrix Φ'. The modified matrix Φ' is symmetric and has the same eigenvectors as the matrix Φ, but its eigenvalues are all positive; thus it is a positive definite symmetric matrix and has a unique inverse Φ'^{-1}.

3.3 *Computer Simulation of the System*

The predictive coding system using adaptive predictors was simulated on a digital computer to determine its effectiveness for coding speech signals. The transmitter and the receiver are illustrated separately in Figs. 4 and 5, respectively. The sampling rate used in this

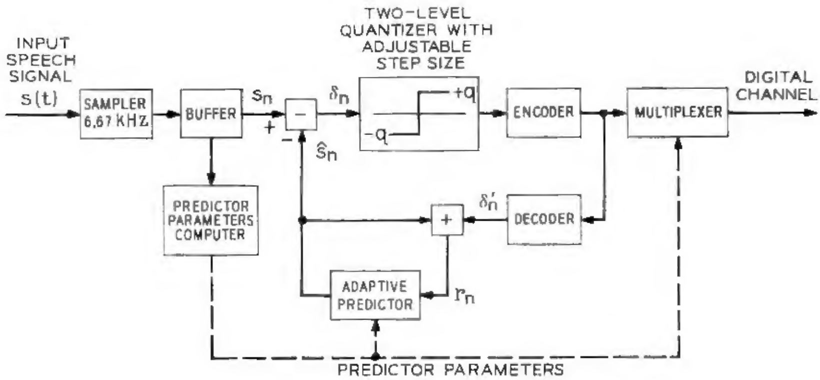

Fig. 4—Transmitter of the predictive coding system.

simulation was 6.67 kHz. Prior to sampling, the input speech signal was filtered with a low-pass filter with 3-dB attenuation at 3.1 kHz and an attenuation of 40 dB or more for frequencies above 3.33 kHz. At the transmitter, the difference δ_n formed by subtracting the predicted value $\hat{s}_n$ from the speech sample s_n was quantized by a *two-level* (1 bit) quantizer with *variable* step size q. The parameter q was readjusted every 5 milliseconds to yield minimum quantization noise power. The parameters of the adaptive predictor were also computed once every 5 milliseconds and sent to the receiver together with the binary difference signal and the step size q of the quantizer. The optimum value of the delay parameter M was obtained by locating the maximum of the correlation coefficient ρ as defined in equation (19) for values of M between 20 and 150. The parameter p was set at 8.

The speech signal was reconstructed at the receiver by a feedback loop containing an adaptive predictor identical to the one used at the transmitter. Here, the predictor too, was reset every 5 milliseconds

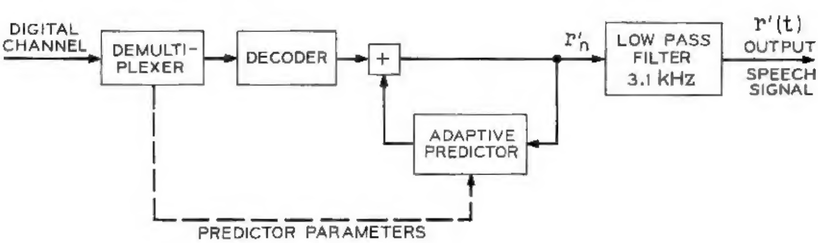

Fig. 5—Receiver of the predictive coding system.

according to the predictor-parameter information received from the transmitter. The reconstructed speech samples were finally smoothed by a 3.1-kHz low-pass filter to form the output speech signal $r'(t)$.

IV. RESULTS OF SUBJECTIVE TESTS

Two different subjective tests were conducted to judge the quality of the reconstructed speech signal produced at the receiver of the predictive coding system. In the first test, trained listeners compared the reconstructed speech signal with speech from a logarithmic PCM (log-PCM) encoder[12] that used the same input signals and a sampling frequency of 6.67 kHz. The compression characteristic employed in a log-PCM encoder is defined by the equation

$$y = \frac{V \log \left[1 + \frac{\mu \mid x \mid}{V}\right]}{\log (1 + \mu)} \operatorname{sgn} x, \tag{26}$$

where y represents the output voltage corresponding to an input signal voltage x, μ is a dimensionless parameter which determines the degree of compression and V is the compressor overload voltage.[12] The compressed signal y was quantized at bit rates varying from 5 bits/sample to 7 bits/sample with $\mu = 100$ and $V = 8 \times$ the rms speech signal voltage.† Speech samples from both male and female speakers were used in these tests. The results of the subjective tests indicated that the quality of the reconstructed speech signal was better than that of log-PCM speech with 5 bits/sample but slightly inferior to one with 6 bits/sample. The corresponding measured signal-to-noise ratios for log-PCM speech were 21 dB and 27 dB, respectively.

In the second test, the reconstructed speech signal was compared with the input speech signal contaminated by additive white noise obtained by randomly inverting the polarity of successive Nyquist samples of the input speech signal.[13] This noise is subjectively similar to the distortion introduced by predictive coding and is therefore particularly appropriate for reproducible comparisons. This noise has an added advantage in that its absolute amplitude at any instant of time is proportional to the absolute amplitude of the input speech signal. This proportionality permits the calculation of a precise signal-to-noise ratio (S/N). Based on the results of these tests, the equivalent S/N of the reconstructed speech in the predictive coding system de-

† The integration time for computing the rms value of the speech signal was several seconds and included speech samples from a number of speakers.

scribed above was found to be about 25 dB which is in good agreement with results obtained by the subjective comparison with log-PCM.

V. ADDITIONAL MODIFICATIONS OF THE PREDICTIVE CODING SYSTEM

5.1 *Spectrum of Quantizing Noise and Its Influence on the Subjective Quality of the Reconstructed Speech*

For frequencies above 500 Hz, the frequency spectrum of voiced speech sounds generally falls off with frequency with an average slope between -6 and -12 dB per octave. The spectrum of quantizing noise in the predictive coding system, on the other hand, is approximately uniform. The signal-to-quantizing noise ratio (S/N) of the reconstructed speech, thus, also falls off with frequency. This is illustrated in Fig. 6 where the spectrum of a short segment of the speech signal is compared with the spectrum of the corresponding quantizing noise. As can be seen, the S/N is very poor at high frequencies. Informal listening tests of the reconstructed speech appeared to confirm the above observation. The quality of the reconstructed speech can thus be improved by a suitable shaping of the spectrum of the quantizing noise so that the S/N is more or less uniform over the entire frequency range of the input speech signal. The desired spectral shaping can be achieved by pre-emphasizing the input speech signal at high frequencies by means of a fixed filter whose amplitude versus frequency characteristic rises with frequency above 500 Hz with a

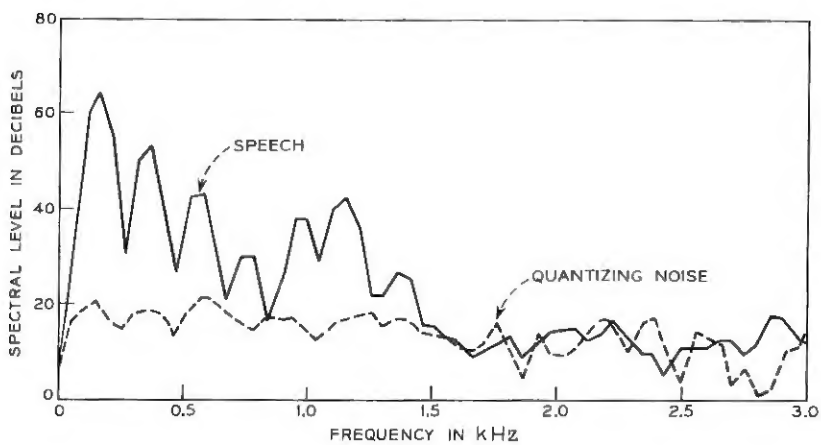

Fig. 6—Spectra of speech and quantizing noise.

slope of 12 dB per octave. The spectral distortion can finally be eliminated by a filter at the output of the receiver whose frequency versus amplitude characteristic is exactly opposite to that of the pre-emphasis filter. The results of computer simulation indicate that the quality of the reconstructed speech in the predictive coding system employing pre-emphasis is considerably better than that of the system without pre-emphasis.

5.2 *Improved Prediction of Voiced Speech*

The redundancy due to the quasi-periodic nature of voiced speech is removed in the predictive coding system described earlier by a predictor $P_1(z)$ consisting of a delay of M samples and an amplifier with gain β as shown in Fig. 3. It is possible to improve the prediction of voiced speech by employing a predictor $P_1(z)$ consisting of two delays and two amplifiers such that

$$P_1(z) = \beta_1 z^{-M} + \beta_2 z^{-2M}. \tag{27}$$

The parameters β_1 and β_2 are calculated by minimizing the mean-square error E_1 defined by

$$E_1 = \langle (s_n - \beta_1 s_{n-M} - \beta_2 s_{n-2M})^2 \rangle_{\text{av}} . \tag{28}$$

The modified predictive coding system including pre-emphasis of the input speech signal together with the second-order predictor $P_1(z)$ as given in equation (27) was simulated on the computer. The results of subjective tests similar to those described in Section IV indicated that the quality of the reconstructed speech was somewhat superior to that of log-PCM speech at 6 bits per sample. The equivalent S/N was found to be 30 dB.

VI. QUANTIZATION OF PREDICTOR PARAMETERS

No attempt was made in the study reported here to quantize the predictor parameters. Preliminary calculations were made to estimate the number of bits required to transmit the information to the receiver. Since the predictor parameters (one delay and nine other coefficients) carry the information about the signal spectrum, it should be possible to encode them at a bit rate comparable to one used in conventional formant vocoders. This suggests a bit rate of approximately 10 kilobits per second for transmitting the binary difference signal (6.67 kb/s) and the predictor parameters (3 kb/s). Recent studies by Kelly[14] indicate that it is indeed possible to encode the transmitted information within 9600 b/s without significant loss in speech quality.

VII. CONCLUSIONS

The study reported here shows that predictive coding is a promising approach to digital encoding of speech signals for high-quality transmission at substantial reductions in bit rate. Unlike past speech coding methods based on the vocoder principle, the predictive coding scheme described here attempts to reproduce accurately the speech *waveform*, rather than its spectrum. Listening tests show that there is only slight, often imperceptible, degradation in the quality of the reproduced speech. Although no detailed investigation of the optimum encoding methods of the predictor parameters was made, preliminary studies suggest that the binary difference signal and the predictor parameters together can be transmitted at bit rates of less than 10 kb/s or several times less than the bit rate required for PCM encoding with comparable speech quality.

REFERENCES

1. Davisson, L. D., "The Theoretical Analysis of Data Compression Systems," Proc. IEEE, *56*, No. 2 (February 1968), pp. 176–186.
2. Cutler, C. C., "Differential Quantization of Communication Signals," U. S. Patent 2-605-361, applied for June 29, 1950; issued July 29, 1952.
3. Oliver, B. N., "Efficient Coding," B.S.T.J., *31*, No. 4 (July 1952), pp. 724–750.
4. Elias, P., "Predictive Coding," IRE Trans. Inform. Theor., *IT-1*, No. 1 (March 1955), pp. 16–33.
5. O'Neal, J. B., Jr., "Predictive Quantizing Systems (Differential Pulse Code Modulation) for the Transmission of Television Signals," B.S.T.J., *45*, No. 5 (May–June 1966), pp. 689–721.
6. Schroeder, M. R., "Vocoders: Analysis and Synthesis of Speech," Proc. IEEE, *54*, No. 5 (May 1966), pp. 720–734.
7. Dudley, H., "Remaking Speech," J. Acoust. Soc. Amer., *11*, No. 2 (October 1939), pp. 169–177.
8. Atal, B. S., and Schroeder, M. R., "Predictive Coding of Speech Signals," Proc. 1967 Conf. on Commun. and Processing, November 1967, pp. 360–361.
9. Atal, B. S., and Schroeder, M. R., "Predictive Coding of Speech Signals," 1968 Wescon Technical Papers, August 1968, paper 8/2.
10. McDonald, R. A., "Signal-to-Noise Performance and Idle Channel Performance of Differential Pulse Code Modulation Systems with Particular Applications to Voice Signals," B.S.T.J., *45*, No. 7 (September 1966), pp. 1123–1151.
11. Faddeev, D. K., and Faddeeva, V. N., *Computational Methods of Linear Algebra*, English Translation by R. C. Williams, San Francisco: W. H. Freeman and Company, 1963, pp. 144–147.
12. Smith, B., "Instantaneous Companding of Quantized Signals," B.S.T.J., *36*, No. 3 (May 1957), pp. 653–709.
13. Schroeder, M. R., "Reference Signal for Signal Quality Studies," J. Acoust. Soc. Amer., *44*, No. 6 (December 1968), pp. 1735–1736.
14. Kelly, J. M., unpublished work.